NATIONAL CENTER FOR EDUCATION STATISTICS

Working Paper Series

The Working Paper Series was created in order to preserve the information contained in these documents and to promote the sharing of valuable work experience and knowledge. However, these documents were prepared under different formats and did not undergo vigorous NCES publication review and editing prior to their inclusion in the series.

U. S. Department of Education
Office of Educational Research and Improvement

NATIONAL CENTER FOR EDUCATION STATISTICS

Working Paper Series

Empirical Evaluation of Social, Psychological, & Educational Construct Variables Used in NCES Surveys

Working Paper No. 95-14 March 1995

Contact: Samuel Peng
Statistical Service and
Methodological Research
(202) 219-1831

U. S. Department of Education
Office of Educational Research and Improvement

U.S. Department of Education
Richard W. Riley
Secretary

Office of Educational Research and Improvement
Sharon P. Robinson
Assistant Secretary

National Center for Education Statistics
Emerson J. Elliott
Commissioner

Data Development Division
Jeanne E. Griffith
Associate Commissioner

National Center for Education Statistics

"The purpose of the Center shall be to collect, analyze, and disseminate statistics and other data related to education in the United States and in other nations."—Section 406(b) of the General Education Provisions Act, as amended (20 U.S.C. 1221e–1).

March 1995

Foreword

Each year a large number of written documents are generated by NCES staff and individuals commissioned by NCES which provide preliminary analyses of survey results and address technical, methodological, and evaluation issues. Even though they are not formally published, these documents reflect a tremendous amount of unique expertise, knowledge, and experience.

The ***Working Paper Series*** was created in order to preserve the valuable information contained in these documents and to promote the sharing of valuable work experience and knowledge. However, these documents were prepared under different formats and did not undergo vigorous NCES publication review and editing prior to their inclusion in the series. Consequently, we encourage users of the series to consult the individual authors for citations.

To receive information about submitting manuscripts or obtaining copies of the series, please contact Suellen Mauchamer at (202) 219-1828 or U.S. Department of Education, Office of Educational Research and Improvement, National Center for Education Statistics, 555 New Jersey Ave., N.W., Room 400, Washington, D.C. 20208-5652.

Susan Ahmed
Acting Associate Commissioner
Statistical Standards and
Methodology Division

Samuel S. Peng
Statistical Service and
Methodological Research

EMPIRICAL EVALUATION OF SOCIAL, PSYCHOLOGICAL, & EDUCATIONAL CONSTRUCT VARIABLES USED IN NCES SURVEYS

March, 1993

Submitted by

Boris Freidlin and Sameena Salvucci

SYNECTICS FOR MANAGEMENT DECISIONS, INC.
3030 Clarendon Boulevard, Suite 305
Arlington, VA 22201

Table of Contents

Preface

This report evaluated composite variables in the National Education Longitudinal Study 1998 (NELS:88) and the Schools and Staffing Survey (SASS). It was prepared by Synectics for Management Decision, Inc., a contractor to the National Center for Education Statistics, as Task 6.2 under Contract No. RN-91-0600.01.

This report was prepared by Boris Freidlin and Sameena Salvucci, research analysts for Synectics. Additional assistance from the Synectics staff and consultants was provided by Arthur Kirsch and Mehrdad Saba working under the direction of Wray Smith, Research Director.

Several key people from National Center for Education Statistics are also worth mentioning. Samuel Peng, Bob Burton, and Steve Kaufman were instrumental in reviewing and providing helpful comments on all drafts. This report would not have been possible without their valuable support.

OVERVIEW

The purpose of this study was to provide an analysis and evaluation of composite variables in National Education Longitudinal Study 1988 (NELS:88) and School and Staffing Survey (SASS) surveys, in a way that will furnish guidance to NCES staff in the more effective use of survey resources. For example, it is reasonable to suppose that more questions are asked in some NCES surveys than are needed for analysis purposes. The study has provided ways of understanding the contributions made by individual survey items through an appraisal of the contributions they make to composite measures of which they are a part. It has also shown how composite variables may provide more reliable measures of the concepts of interest than do individual survey items, and how they may permit the more effective summarization and communication of survey results.

This study uses the term "composite" to denote the use of multiple survey items for a single measure. The composite variables that were examined include the following:

(1) Self concept and locus of control
(2) Socioeconomic status (SES)
(3) School climate

Additional variables designated by NCES and variables found to be of interest during the course of the study were also analyzed. The guiding question in the study analysis was whether a particular variable contributed to or detracted from the efficacy of a composite variable. This question was approached by inquiring about whether to include the variable in the factor analysis. The initial consideration was whether the variable conceptually fit with the other components of the composite.

The strategy followed in this study employed several factor analyses followed by calculation of reliability and validity estimates on a given composite variable. While the field costs of surveys have escalated, the computer has made analysis comparatively quick and cheap, making it cost effective to conduct a variety of analyses of survey items and their composites. A variety of factor analyses were conducted, to see what relationships are stable across several analyses.

A summary of the analyses follows.

I. Self concept and Locus of control

The following group of items were identified as the most efficient self-concept and locus of control composites.

Self-concept

"I feel good about myself"
"I'm a person of worth, equal of others"
"I am able to do things as well as others"
"On the whole I am satisfied with myself"
"When I make plans I can make them work"

Locus of control

"Good luck is more important than hard work"
"Every time I get ahead something stops me"
"Plans hardly work out, makes me unhappy"
"I don't have enough control over my life"
"Chance and luck important in my life"

It is important to note that the three remaining items dropped out of the self-concept composite:

"I certainly feel useless at times"
"At time I feel I am no good at all"
"I feel I do not have much to be proud of"

Also, the item "When I make plans I can make them work" typically had been part of the locus of control composite. This suggests that the analysis could possibly have been influenced by the difference in response to the reversed score items and possible failure by the respondents to recognize the repeated shifts to and from reversed score questions. Even though reversed items were used to avoid "response set", this technique added to respondents confusion. Perhaps these items should be tried out with consistent direction (all positive or negative) or at least arranged in two separate groups.

II. Socio-Economic Status

Evaluation of the SES composite showed that a comparably valid and reliable SES composite could be constructed from the following items:

Father's education
Mother's education
Family income (household items list if income missing).

This composite differed from currently used SES composite in two ways:

a) The current composite used data from the parent file and only if all the items were missing in parent file the data were taken from the student file. The proposed composite used the student file information each time an item was missing in the parent file.

b) The current composite used parent's occupation data items which were difficult to recode. The proposed composite did not use parents occupation data.

The analysis indicated that not only was it easier to calculate the proposed SES composite, its validity and reliability were equal to those of the currently used SES composite.

III. School Climate

Two "school climate" item pools (one from the student file and one from the school file) were grouped into "school climate" composites. Unfortunately, the grouping resulted in a dramatic drop of the predictive power, reducing usefulness of the composites.

IV. Overall predictive power

Predictive power of SES, locus of control, self-concept and school climate items combined was evaluated. The model indicated that

a) Only the SES composite was capable of condensing the information of the original items while preserving the predictive power.

b) The locus of control composite lost about 15% of its R^2 compared to individual items, but was still a relatively good predictor of the achievement scores.

c) The rest of the composites/items were of almost no use as predictors of student achievement scores.

V. SASS composites

Separate analysis was done on the two SASS item pools to investigate potential "school climate" and "perception of school problem" composites. Adequate groupings were achieved for both item pools (relatively good fit and high reliability). Unfortunately, SASS files did not contain any score variables, so no examination of predictive power was possible for the SASS composites.

VI. Summary and Findings

The report identified some ways of simplifying and optimizing the existing composites. At the same time it confirmed the statistical foundation of the SES, Locus of control and Self concept composites. No strong "school climate" composites surfaced.

The study results offer promise for revision of survey instrument contents to help shorten surveys, reduce response burden, heighten response rates, improve communications with data users, and bring about increased reliability of measurement.

National Education Longitudinal Study: 1988 Base Year (NELS:88)

A. SELF-CONCEPT AND LOCUS OF CONTROL COMPOSITES ANALYSIS

I. Introduction

This analysis evaluated self-concept and locus of control composites currently in the National Education Longitudinal Study of 1988 (NELS:88). It looked into the correctness of the grouping of the original components into composites, and their reliability and validity (predictive power).

The NELS:88 used 13 variables to construct self-concept and locus of control composites. For the purposes of this analysis the following two versions of each composite were constructed:

a) A version comparable with High School and Beyond (HS&B) and National Longitudinal Study NLS-72 (short version, uses fewer variables).

b) A full version using all the available variables.

Self concept and locus of control items were all in student question #44. The values of these items range from 1 to 4, meaning "strongly agree", 4 "strongly disagree".

SELF-CONCEPT:

version1	version2	label
BYS44A	BYS44A	I feel good about myself
BYS44D	BYS44D	I'm a person of worth, equal of others
BYS44E	BYS44E	I am able to do things as well as others
BYS44H	BYS44H	On the whole, I am satisfied with myself
	BYS44I	I certainly feel useless at times
	BYS44J	At times I feel I am no good at all
	BYS44L	I feel I do not have much to be proud of

The scores for the four first items were reversed. Each of the items was standardized (mean=0 and std=1) and all nonmissing components averaged. The observations with all the items missing were assigned missing values.

LOCUS OF CONTROL:

version1	version2	label
BYS44C	BYS44C	Good luck is more important than hard work
BYS44F	BYS44F	Every time I get ahead something stops me
BYS44G	BYS44G	Plans rarely work out, makes me unhappy
	BYS44B	I don't have enough control over my life
	BYS44K	When I make plans I can make them work
	BYS44M	Chance and luck important in my life

The scores for BYS44K were reversed. Each of the items was standardized (mean=0 std=1) and all nonmissing components were averaged. The observations with all the items missing were assigned missing values.

II. Analysis Plan

First, it would be helpful to give definition and interpretation to a number of coefficients used in this analysis:

1) h^2 - communality, which is listed for each of the items in each of the factor analysis solutions. Communality is the portion of the item's variance accounted by all common factor. h^2 is calculated as the sum of the squared factor loadings.

2) Root Mean Square Off-diagonal Partials (RMS) represent the partial correlation among the items after removing effects of the common factors. The assumption of the common factor model implies that RMS should be 0. Therefore, RMS is a good way to assess goodness-of-fit of the model: the closer RMS is to 0 the better. RMS is calculated as the squared root of the mean of the off diagonal squared partial correlations.

3) Cronbach's Alpha is a measure of reliability. It is defined as the portion of the composite's total variance that is attributable to a common source. Cronbach's Alpha is calculated as follows:

$$\alpha = \frac{k}{(k-1)}\left(1 - \frac{\Sigma \sigma_i^2}{\sigma^2}\right)$$

k is the number of items in the composite

σ_i^2 *is the sum of item variances*

σ^2 *is the sum of item variances and covariances*

Various factor analysis techniques were used to check the grouping of the above component variables into self concept and locus of control composites. As a result of using listwise deletion in these analyses, 22605 observations with nonmissing data were used. All analyses were conducted without applying sample weights.

Two different factor analysis methods (available in SAS PROC FACTOR) were used:

a) *Principal Factor Analysis*, where prior communality estimate for each item was set to squared multiple correlation.

b) *Maximum Likelihood Factor Analysis*, where prior communality estimate for each item was set to squared multiple correlation.

Each of the factor analyses included:

- Varimax rotation
- Factor loadings, communalities, portion variance explained by each factor and Root Mean Square Off-diagonal partials were corded for each run.

Loadings greater than .4 were considered to be meaningful. Suggested composite groupings were identified and the composites were calculated using the same procedures as described above.

The sample was randomly split into two subsamples and factor analysis was performed on each of the subsamples. Comparison of the two subsample solutions and the solution obtained from the complete sample was made in order to measure the stability of the final results.

Reliability of the new suggested composites as well as the old ones was evaluated by Cronbach's Alpha using SPSS RELIABILITY Procedure.

Validity (Predictive Power) of the composites was measured by the correlations between the composites and Standardized Math, Science and Reading scores taken from the NELS:88.

Stepwise regression for each of the test scores using individual component items for each of the composites as independent items was done to compare the composite's predictive power with that of the component items.

III. Results

FACTOR ANALYSIS (TWO FACTORS):

Principal Factor Analysis

Proportion variance explained=.344 RMS=.079

Rotated Factor Pattern

variables	F1	F2	h^2	label
BYS44A	.703	.107	.505	I feel good about myself
BYS44H	.683	.180	.499	On the whole, I am satisfied with myself
BYS44D	.561	.122	.329	I'm a person of worth, equal of others
BYS44E	.512	.093	.271	I am able to do things as well as others
BYS44K	.452	.204	.246	When I make plans I can make them work
BYS44G	.306	.553	.400	Plans hardly work out, makes me unhappy
BYS44F	.236	.546	.353	Every time I get ahead something stops me
BYS44M	-.006	.515	.265	Chance and Luck important in my life
BYS44C	.031	.494	.245	Good luck is more important than hard work
BYS44B	.226	.470	.272	I don't have enough control over my life
BYS44L	.417	.457	.383	I feel I do not have much to be proud of
BYS44J	.432	.433	.375	At times I feel I am no good at all
BYS44I	.385	.427	.330	I certainly feel useless at times
Proportion of variance explained.	.189	.155		

Maximum Likelihood

Proportion variance explained=.344
RMS=.078

Rotated Factor Pattern

variables	F1	F2	h^2	label
BYS44A	.718	.118	.529	I feel good about myself
BYS44H	.697	.187	.521	On the whole, I am satisfied with myself
BYS44D	.556	.128	.326	I'm a person of worth, equal of others
BYS44E	.500	.107	.261	I am able to do things as well as others
BYS44K	.431	.233	.241	When I make plans I can make them work
BYS44G	.274	.576	.407	Plans hardly work out, makes me unhappy
BYS44F	.209	.568	.367	Every time I get ahead something stops me
BYS44M	.036	.477	.227	Chance and Luck important in my life
BYS44B	.216	.472	.270	I don't have enough control over my life
BYS44J	.401	.470	.382	At times I feel I am no good at all
BYS44I	.354	.467	.344	I certainly feel useless at times
BYS44L	.405	.466	.381	I feel I do not have much to be proud of
BYS44C	.040	.456	.209	Good luck is more important than hard work
Proportion of variance explained	.182	.162		

FACTOR ANALYSIS (THREE FACTORS):

Principal Factor Analysis

Proportion variance explained = .401
RMS = .05

Rotated Factor Pattern

variables	F1	F2	F3	h^2	label
BYS44H	.662	.218	.132	.504	On the whole, I am satisfied with myself
BYS44A	.646	.264	.030	.488	I feel good about myself
BYS44D	.605	.063	.138	.389	I'm a person of worth, equal of others
BYS44E	.542	.068	.099	.308	I am able to do things as well as others
BYS44K	.433	.194	.147	.247	When I make plans I can make them work
BYS44L	.386	.294	.377	.378	I feel I do not have much to be proud of
BYS44I	.187	.744	.152	.611	I certainly feel useless at times
BYS44J	.253	.711	.167	.598	At times I feel I am no good at all
BYS44C	.076	.006	.611	.379	Good luck is more important than hard work
BYS44M	.019	.076	.568	.328	Chance and Luck important in my life
BYS44G	.277	.315	.454	.383	Plans hardly work out, makes me unhappy
BYS44F	.192	.339	.428	.335	Every time I get ahead something stops me
BYS44B	.206	.246	.403	.265	I don't have enough control over my life
Proportion of variance explained	.163	.122	.116		

Maximum Likelihood

Proportion of variance explained = .401
RMS = .048

Rotated Factor Pattern

variables	F1	F2	F3	h^2	label
BYS44H	.689	.138	.185	.528	On the whole, I am satisfied with myself
BYS44A	.675	.035	.240	.515	I feel good about myself
BYS44D	.575	.142	.073	.356	I'm a person of worth, equal of others
BYS44E	.514	.107	.074	.281	I am able to do things as well as others
BYS44K	.434	.172	.160	.243	When I make plans I can make them work
BYS44L	.396	.396	.258	.380	I feel I do not have much to be proud of
BYS44C	.072	.587	.004	.350	Good luck is more important than hard work
BYS44M	.018	.559	.072	.318	Chance and Luck important in my life
BYS44G	.281	.485	.258	.381	Plans hardly work out, makes me unhappy
BYS44F	.202	.455	.290	.332	Every time I get ahead something stops me
BYS44B	.216	.420	.209	.267	I don't have enough control over my life
BYS44I	.194	.177	.763	.651	I certainly feel useless at times
BYS44J	.264	.196	.710	.613	At times I feel I am no good at all
Proportion of variance explained	.165	.122	.114		

RELIABILITY:

Cronbach's Alpha was calculated for the following suggested group of items:

SC1: BYS44H, BYS44A, BYS44D, BYS44E
STD Alpha=.735

SC2: BYS44H, BYS44A, BYS44D, BYS44E, BYS44L, BYS44J, BYS44I
STD Alpha=.787

SC3: BYS44H, BYS44A, BYS44D, BYS44E, BYS44K
STD Alpha=.744

LC1: BYS44C, BYS44G, BYS44F
STD Alpha=.575

LC2: BYS44C, BYS44M, BYS44G, BYS44F, BYS44B, BYS44K,
STD Alpha=.680

LC3: BYS44C, BYS44M, BYS44G, BYS44F, BYS44B
STD Alpha=.680

UL: BYS44I BYS44J
STD Alpha=.776

VALIDITY:

Squared correlations between different groups of variables and math, science and reading scores:

composite	math	science	reading
SC1 BYS44H,BYS44A,BYS44D,BYS44E	.006	.006	.005
SC2 BYS44H,BYS44A,BYS44D,BYS44E,BYS44L,BYS44I,BYS44J	.019	.018	.017
SC3 BYS44H,BYS44A,BYS44D,BYS44E,BYS44K	.005	.006	.005
LC1 BYS44C,BYS44G,BYS44F	.082	.068	.093
LC2 BYS44C,BYS44M,BYS44G,BYS44F,BYS44B,BYS44K	.085	.074	.095
LC3 BYS44C,BYS44M,BYS44G,BYS44F,BYS44B	.103	.088	.115
UL BYS44I,BYS44J	.016	.016	.011

Stepwise regressions were performed using standardized **math, science** and **reading** scores as a dependent variable and self-concept and locus of control items as independent variables. At significance level for staying (sls) =.05 and significance level of entry (sle) =.05 the following variables were selected:

SC1

Dependent variable math score

step	variable	label	model R^2
1	BYS44H	On the whole, I am satisfied with myself	.008
2	BYS44A	I feel good about myself	.015
3	BYS44D	I'm a person of worth, equal of others	.021
4	BYS44E	I am able to do things as well as others	.023

Dependent variable science score

step	variable	label	model R^2
1	BYS44D	I'm a person of worth, equal of others	.009
2	BYS44A	I feel good about myself	.012
3	BYS44H	On the whole, I am satisfied with myself	.018
4	BYS44E	I am able to do things as well as others	.021

Dependent variable reading score

step	variable	label	model R^2
1	BYS44D	I'm a person of worth, equal of others	.011
2	BYS44A	I feel good about myself	.019
3	BYS44H	On the whole, I am satisfied with myself	.026
4	BYS44E	I am able to do things as well as others	.028

SC2

Dependent variable math score

step	variable	label	model R^2
1	BYS44L	I feel I do not have much to be proud of	.031
2	BYS44A	I feel good about myself	.039
3	BYS44J	At times I feel I am no good at all	.047
4	BYS44H	On the whole, I am satisfied with myself	.052
5	BYS44D	I'm a person of worth, equal of others	.054
6	BYS44E	I am able to do things as well as others	.055
7	BYS44I	I certainly feel useless at times	.056

Dependent variable science score

step	variable	label	model R^2
1	BYS44L	I feel I do not have much to be proud of	.027
2	BYS44J	At times I feel I am no good at all	.033
3	BYS44A	I feel good about myself	.044
4	BYS44D	I'm a person of worth, equal of others	.048
5	BYS44H	On the whole, I am satisfied with myself	.050
6	BYS44E	I am able to do things as well as others	.051

BYS44I not entered at significance level=.05

Dependent variable reading score

step	variable	label	model R^2
1	BYS44L	I feel I do not have much to be proud of	.037
2	BYS44A	I feel good about myself	.049
3	BYS44D	I'm a person of worth, equal of others	.058
4	BYS44J	At times I feel I am no good at all	.062
5	BYS44H	On the whole, I am satisfied with myself	.064
6	BYS44E	I am able to do things as well as others	.064

BYS44I not entered at significance level=.05

SC3

Dependent variable math score

step	variable	label	model R^2
1	BYS44H	On the whole, I am satisfied with myself	.008
2	BYS44A	I feel good about myself	.015
3	BYS44D	I'm a person of worth, equal of others	.021
4	BYS44E	I am able to do things as well as others	.023

BYS44K not entered at significance level=.05

Dependent variable science score

step	variable	label	model R^2
1	BYS44D	I'm a person of worth, equal of others	.009
2	BYS44A	I feel good about myself	.012
3	BYS44H	On the whole, I am satisfied with myself	.018
4	BYS44E	I am able to do things as well as others	.021

BYS44K not entered at significance level=.05

Dependent variable reading score

step	variable	label	model R^2
1	BYS44D	I'm a person of worth, equal of others	.011
2	BYS44A	I feel good about myself	.019
3	BYS44H	On the whole, I am satisfied with myself	.026
4	BYS44E	I am able to do things as well as others	.028

BYS44K not entered at significance level=.05

LC1

Dependent variable math score

step	variable	label	model R^2
1	BYS44C	Good luck is more important than hard work	.046
2	BYS44F	Every time I get ahead, something stops me	.073
3	BYS44G	Plans hardly work out, makes me unhappy	.083

Dependent variable science score

step	variable	label	model R^2
1	BYS44C	Good luck is more important than hard work	.041
2	BYS44G	Plans hardly work out, makes me unhappy	.062
3	BYS44F	Every time I get ahead, something stops me	.069

Dependent variable reading score

step	variable	label	model R^2
1	BYS44C	Good luck is more important than hard work	.062
2	BYS44F	Every time I get ahead, something stops me	.087
3	BYS44G	Plans hardly work out, makes me unhappy	.097

LC2

Dependent variable math score

step	variable	label	model R^2
1	BYS44M	Chance and luck important in my life	.085
2	BYS44G	Plans hardly work out, makes me unhappy	.105
3	BYS44F	Every time I get ahead, something stops me	.113
4	BYS44C	Good luck is more important than hard work	.119
5	BYS44K	When I make plans I can make them work	.122

BYS44B not entered at significance level=.05

Dependent variable science score

step	variable	label	model R^2
1	BYS44M	Chance and luck important in my life	.080
2	BYS44G	Plans hardly work out, makes me unhappy	.096
3	BYS44C	Good luck is more important than hard work	.102
4	BYS44F	Every time I get ahead, something stops me	.105
5	BYS44K	When I make plans I can make them work	.107

BYS44B not entered at significance level=.05

Dependent variable reading score

step	variable	label	model R^2
1	BYS44M	Chance and luck important in my life	.096
2	BYS44G	Plans hardly work out, makes me unhappy	.116
3	BYS44C	Good luck is more important than hard work	.128
4	BYS44F	Every time I get ahead, something stops me	.135
5	BYS44K	When I make plans I can make them work	.138

BYS44B not entered at significance level=.05

LC3

Dependent variable math score

step	variable	label	model R^2
1	BYS44M	Chance and luck important in my life	.085
2	BYS44G	Plans hardly work out, makes me unhappy	.105
3	BYS44C	Good luck is more important than hard work	.113
4	BYS44F	Every time I get ahead, something stops me	.118

BYS44B not entered at significance level=.05

Dependent variable science score

step	variable	label	model R^2
1	BYS44M	Chance and luck important in my life	.080
2	BYS44G	Plans hardly work out, makes me unhappy	.096
3	BYS44C	Good luck is more important than hard work	.101
4	BYS44F	Every time I get ahead, something stops me	.105

BYS44B not entered at significance level=.05

Dependent variable reading score

step	variable	label	model R^2
1	BYS44M	Chance and luck important in my life	.096
2	BYS44G	Plans hardly work out, makes me unhappy	.116
3	BYS44C	Good luck is more important than hard work	.128
4	BYS44F	Every time I get ahead, something stops me	.135

BYS44B not entered at significance level=.05

Stepwise regressions were performed using standardized **math, science** and **reading** scores as a dependent variable and all 13 individual items as independent variables. At significance level for staying (sls) =.05 and significance level of entry (sle) =.05 the following variables were selected:

Dependent variable math score

step	variable	label	model R^2
1	BYS44M	Chance and luck important in my life	.085
2	BYS44G	Plans hardly work out, makes me unhappy	.105
3	BYS44F	Every time I get ahead, something stops me	.113
4	BYS44A	I feel good about myself	.120
5	BYS44L	I feel I do not have much to be proud of	.126
6	BYS44C	Good luck is more important than hard work	.131
7	BYS44K	When I make plans I can make them work	.133
8	BYS44H	On the whole, I am satisfied with myself	.135
9	BYS44D	I'm a person of worth, equal of others	.137
10	BYS44E	I am able to do things as well as others	.137
11	BYS44J	At times I feel I am no good at all	.138

BYS44B and BYS44I not entered at significance level=.05

Dependent variable science score

step	variable	label	model R^2
1	BYS44M	Chance and luck important in my life	.080
2	BYS44G	Plans hardly work out, makes me unhappy	.096
3	BYS44C	Good luck is more important than hard work	.102
4	BYS44A	I feel good about myself	.106
5	BYS44L	I feel I do not have much to be proud of	.111
6	BYS44F	Every time I get ahead, something stops me	.115
7	BYS44D	I'm a person of worth, equal of others	.117
8	BYS44J	At times I feel I am no good at all	.120
9	BYS44K	When I make plans I can make them work	.121
10	BYS44E	I am able to do things as well as others	.123
11	BYS44I	I certainly feel useless at times	.123
12	BYS44H	On the whole, I am satisfied with myself	.124

BYS44B not entered at significance level=.05

Dependent variable reading score

step	variable	label	model R^2
1	BYS44M	Chance and luck important in my life	.096
2	BYS44G	Plans hardly work out, makes me unhappy	.116
3	BYS44C	Good luck is more important than hard work	.128
4	BYS44A	I feel good about myself	.137
5	BYS44L	I feel I do not have much to be proud of	.146
6	BYS44F	Every time I get ahead, something stops me	.152
7	BYS44D	I'm a person of worth, equal of others	.156
8	BYS44K	When I make plans I can make them work	.158
9	BYS44H	On the whole, I am satisfied with myself	.159
10	BYS44I	I certainly feel useless at times	.160
11	BYS44J	At times I feel I am no good at all	.160
12	BYS44E	I am able to do things as well as others	.161

BYS44B not entered at significance level=.05

Stepwise regressions were performed using standardized **math, science** and **reading** scores as a dependent variable and LC3, SC3, BYS44L and UL as independent variables. At significance level for staying (sls) =.05 and significance level of entry (sle) =.05 the following variables were selected:

Dependent variable math score

step	variable	label	model R^2
1	LC3		.103
2	SC3		.105
3	BYS44L	I feel I do not have much to be proud of	.108

UL not entered at significance level=.05

Dependent variable science score

step	variable	label	model R^2
1	LC3		.088
2	SC3		.090
3	BYS44L	I feel I do not have much to be proud of	.092

UL not entered at significance level=.05

Dependent variable reading score

step	variable	label	model R^2
1	LC3		.115
2	SC3		.119
3	BYS44L	I feel I do not have much to be proud of	.123
4	UL		.124

IV. Analysis

For the two-factor model both Principal Factor and Maximum Likelihood analyses produced similar results with only a modest proportion of total variance explained (.344) and relatively high RMS (.079 for Principal Factor and .078 for Maximum Likelihood). A number of variables (BYS44I, BYS44J, BYS44L) had high loadings on both factors, suggesting that they were either complex variables or were forced into the two factors thus making their interpretation difficult. The two-factor analysis results indicated that more than two factors should be used in the model.

For the three factor model Principal Factor and Maximum Likelihood Factor analyses second and third factors were interchanged; apart from that the results were similar. Both methods gave better fit than the two factor model: proportion of variance explained = .401 and RMS under .05. Each of the variables, except BYS44L, was highly loaded on one factor only, suggesting following grouping of items:

SC3: Self-concept

BYS44A	I feel good about myself
BYS44D	I'm a person of worth, equal of others
BYS44E	I am able to do things as well as others
BYS44H	On the whole, I am satisfied with myself
BYS44K	When I make plans I can make them work

LC3: Locus of control

BYS44C	Good luck is more important than hard work
BYS44F	Every time I get ahead something stops me
BYS44G	Plans hardly work out, makes me unhappy
BYS44B	I don't have enough control over my life
BYS44M	Chance and Luck important in my life

with

UL:	BYS44I	I certainly feel useless at times
	BYS44J	At times I feel I am no good at all

being a separate factor. BYS44L had moderate to low loadings (below our .4 cutoff) on all 3 factors.

It should be pointed out that BYS44A, BYS44D, BYS44E, BYS44H, BYS44K are the five reversed score questions and they always grouped together (in the two-factor and the three-factor solutions), even though they seemed to belong to the different factors.

Maximum Likelihood Factor Analysis on the two random subsets of the whole sample gave results almost identical to those above, confirming stability of the solution.

Reliability of the each composite was estimated by Cronbach's Alpha which can be interpreted as the squared correlation between the specific composite score a person obtains and the score he or she would have obtained if **all** possible items were used.

Cronbach's Alpha for SC2: BYS44A, BYS44D, BYS44E, BYS44H, BYS44I, BYS44J, BYS44L was highest among the self-concept composites. Adding BYS44K to SC1 slightly increased reliability. Even though reliability is increased by adding BYS44I, BYS44J, BYS44l to the self-concept composite, the increase is rather low considering that the composite went from four items to seven items.

Dropping BYS44K from LC2 did not make any significant difference in terms of Cronbach's Alpha. However, addition of BYS44M to LC1 significantly improved reliability, thus making LC3: BYS44C, BYS44F, BYS44G, BYS44B, BYS44M a "better" composite than LC1 or LC2.

The correlations between the composites and math,science and reading scores indicated that composite SC2: BYS44A, BYS44D, BYS44E, BYS44H, BYS44I, BYS44J, BYS44L had much higher predictive power than the two other self-concept composites. It must be pointed out, though, that individually, the component items were much better predictors than self-concept composites: using reading score BYS44L alone had R^2=.037 (with full model R^2=.064) while SC2 had R^2=.017

In case of locus of control composites LC3: BYS44C, BYS44F, BYS44G, BYS44B, BYS44M did a better job than the two other composites: LC3 squared correlations with test scores were the highest. Individual items were better predictors than the composites.

Moreover, for each achievement score SC3 and LC3 combined had lower model R^2 than BYS44M, BYS44G and BYS44C (BYS44F instead of BYS44C for math score). Again, individual items had higher predictive power than composites.

V. Conclusion

Suggested composites are

Self-concept

BYS44A	I feel good about myself
BYS44D	I'm a person of worth, equal of others
BYS44E	I am able to do things as well as others
BYS44H	On the whole, I am satisfied with myself
BYS44K	When I make plans I can make them work

Locus of control

BYS44C	Good luck is more important than hard work
BYS44F	Every time I get ahead something stops me
BYS44G	Plans hardly work out, makes me unhappy
BYS44B	I don't have enough control over my life
BYS44M	Chance and Luck important in my life

The results suggested that the analysis might have been influenced by the difference in the response to the reversed score items and possible failure by the respondents to recognize the repeated shifts to and from the reversed score questions:

a) The reversed score items always grouped together, thus making the presence of BYS44k in self-concept questionable (it might be only due to the inflated correlation among consistent direction questions).

b) BYS44I, BYS44J, BYS44L (which dropped out of self-concept) came directly after reversed score question BYS44H and were split by reverse score question BYS44K.

Even though psychometricians have used reversed items to avoid "response set", in the minds of the respondents it adds confusion and perhaps these items should be tried out with consistent direction (all positive or negative) or at least arranged in two separate groups.

Individual items used in a stepwise regression produced significantly better prediction for the achievement scores than the composites: thus it is suggested that the separate items should be used if predictive power is the most important goal.

National Education Longitudinal Study: 1988 Base Year (NELS:88)

B. SOCIO-ECONOMIC STATUS (SES) COMPOSITES ANALYSIS

I. Introduction

This analysis evaluated the SES composite used in NELS in terms of the method of construction, items used in the composite construction, predictive power of the composites (validity) and reliability (for the new composites). It specifically looked into the possibility of constructing a new SES composite of comparable validity but using fewer or less complex items.

NELS:88 SES composite (**BYSES**) was constructed in the following way:

a) Parent questionnaire items:

Father's education level
Mother's education level
Father's occupation (recoded using Duncan SEI Scale)
Mother's occupation (recoded using Duncan SEI Scale)
Family Income

Each item was standardized (mean=0 std=1) and all nonmissing components averaged.

b) In the case where **all** the parent items were missing (8.1 percent) student data were used:

Father's education level
Mother's education level
Father's occupation (recoded using Duncan SEI Scale)
Mother's occupation (recoded using Duncan DEI Scale)
Household items list

Household items list was used instead of income: if more than seven household questions were answered they were averaged. Each item was standardized (mean=0 std=1) and all nonmissing components averaged.

II. Analysis Plan

Three new SES composites were constructed:

1) **SES1:**

Father's education from parent file; if missing, student file value was used.
Mother's education from parent file; if missing, student file value was used.
Family income from parent file.

2) **SES2:**

Father's education from parent file; if missing, student file value was used.
Mother's education from parent file; if missing, student file value was used.
Family income from parent file; if missing, household items list from student file was used.

3) **SES3:**

Father's education from parent file; if missing, student file value was used.
Mother's education from parent file; if missing, student file value was used.
Family income from parent file.
Family composition: 1 if two parents 0 otherwise.

To measure predictive power of a SES composite math, science, reading and history standardized achievement scores from base NELS:88 file were used. Correlation coefficients between the four SES composites (BYSES, SES1, SES2, SES3) and the four achievement scores were computed using listwise deletion.

The correlation coefficient computations were repeated with sample stratified by race to look for possible differences in SES definition in the strata.

Reliability of the new composites was evaluated by Cronbach's Alpha.

Stepwise regressions were performed using the achievement scores as dependent variables and composite items as independent variables (SLS=.05 SLE=.05) to see how composite predictive power compares to that of the individual items.

III. Results

SES Composite	No. of nonmissing observations
BYSES	24588
SES1	22631
SES2	24530
SES3	24481

Predictive Power:

Correlation coefficients were computed using listwise deletion resulting in 21647 observations used:

SES composite	science score	math score	history score	reading score
BYSES	.38948	.44358	.40951	.40504
SES1	.39504	.44852	.41145	.40311
SES2	.39680	.44998	.41150	.40465
SES3	.38167	.43063	.39149	.38451

Correlation coefficients stratified by race were computed using listwise deletion resulting in:

1) Strata: Race=Asian 1306 observations used:

SES composite	science score	math score	history score	reading score
BYSES	.37824	.42787	.44148	.44643
SES1	.38097	.42875	.44025	.45233
SES2	.38162	.43016	.44157	.45466
SES3	.37402	.42444	.42973	.44197

2) Strata: Race=Hispanic 2588 observations used:

SES composite	science score	math score	history score	reading score
BYSES	.26576	.33225	.31164	.31898
SES1	.26815	.32600	.31118	.31565
SES2	.26874	.32509	.30994	.31481
SES3	.26105	.31765	.29157	.29460

3) Strata: Race=Black 2552 observations used:

SES composite	science score	math score	history score	reading score
BYSES	.28568	.30755	.29296	.31128
SES1	.29512	.31865	.30665	.30308
SES2	.29452	.31684	.30369	.30279
SES3	.26925	.27987	.26436	.27099

4) Strata: Race=White 14771 observations used:

SES composite	science score	math score	history score	reading score
BYSES	.34113	.40321	.37072	.35682
SES1	.35203	.41315	.37465	.35862
SES2	.35305	.41416	.37407	.35951
SES3	.32789	.38515	.34884	.33250

Reliability:

Cronbach's Alpha was calculated for following groups of items:

SES composite	Cronbach's Alpha
SES1 (father's education, mother's education, family income)	.738782
SES2 (father's education, mother's education,family income/household items)	.736212
SES3 (father's education, mother's education, family income, family composition)	.664295

Stepwise regressions were performed using science, math, history and reading scores as a dependent variable and father's education, mother's education, family income, household items standardized list and family configuration as independent variables. AS a squared correlation coefficients of the SES composites and achievement scores on the same subsamples were calculated for comparison:

Dependent variable science score (n=19128)

step	variable	model R^2
1	mother's education	.1012
2	Family income	.1355
3	Fathers's education	.1477
4	Household items	.1528
5	Family composite	.1530

SES composite	squared correlation coefficient
BYSES	.1430
SES1	.1472
SES2	.1472
SES3	.1364

Dependent variable math score (n=19141)

step	variable	model R^2
1	mother's education	.1377
2	Family income	.1812
3	Father's education	.1939
4	Household items	.1990
5	Family composite	.1992

SES composite	squared correlation coefficient
BYSES	.1896
SES1	.1926
SES2	.1926
SES3	.1782

Dependent variable history score (n=19056)

step	variable	model R^2
1	mother's education	.1093
2	Family income	.1461
3	Fathers's education	.1602
4	Household items	.1658

family composite not entered at significance level=.05

SES composite	squared correlation coefficient
BYSES	.1591
SES1	.1598
SES2	.1598
SES3	.1455

Dependent variable reading score (n=19150)

step	variable	model R^2
1	Mother's education	.1065
2	Family income	.1411
3	Fathers's education	.1560
4	Household items	.1591

family composite not entered at significance level=.05

SES composite	squared correlation coefficient
BYSES	.1567
SES1	.1558
SES2	.1558
SES3	.1430

IV. Analysis

SES2 had the highest correlation with science, math and history scores while BYSES had the highest correlation with the reading score. At the same time the correlations between BYSES, SES1 and SES2 and the achievement scores were so close (maximum difference <1.5%) that none of the composites could be deemed the best on the predictive power merit only.

One of the more desirable properties of the SES composite is simplicity of calculation. Both SES1 and SES2 do not involve data collection on parents occupation and do not require any recoding and therefore have clear advantages over using BYSES. In addition SES1 and SES2 use available data items more efficiently than BYSES by substituting student items for missing parent items when possible.

Another issue that should be addressed is the number of students for whom the composite items are available. The most important items in the composites come from the parents questionnaire. Eight percent of the students did not have any parent questionnaire information available. As a result approximately seven percent of the students had the SES2 composite based on household items only. At the same time the BYSES composite utilized parent occupation information. That explains why the correlation with achievement scores using pairwise deletion is slightly higher for BYSES than for SES2 with a maximum difference of <3.4%. Reliability of SES1 and SES2 composites were essentially the same; consequently SES2 seemed to be the most efficient SES composite.

Correlation coefficients in the sample stratified by race showed the correlations for Hispanic and Black students were lower than the correlations for the overall sample. This may indicate that the SES composite for these two groups might be constructed or interpreted separately.

In addition, stepwise regressions indicated that the composites BYSES, SES1 and SES2 were almost as good predictors as the individual items.

V. Conclusion

On the basis of predictive power, simplicity of calculation and availability

SES2:

Father's education from parent file, if missing student file value was used.
Mother's education from parent file, if missing student file value was used.
Family income from parent file, if missing household items list from student file was used.

seemed to be the best choice for the SES composite in NELS:88 file.

C. SCHOOL CLIMATE (SCHOOL LEVEL) COMPOSITES ANALYSIS

I. Introduction

This section explores possibilities of constructing school climate composites. The analysis was performed at the school level, using the following 63 items from NELS:88 Base Year School File:

variable	label
BYSC47A	conflict between teachers and administrators
BYSC47B	discipline is emphasized at this school
BYSC47C	students place a priority on learning
BYSC47D	classroom environment is structured
BYSC47E	teachers encourage students to do their best
BYSC47F	students are expected to do homework
BYSC47G	teachers morale is high
BYSC47H	teachers have negative attitude about students
BYSC47I	teachers have difficulty motivating student
BYSC47J	school day for students is structured
BYSC47K	deviation from school rules not tolerated
BYSC47L	school environment is flexible
BYSC47M	teachers respond to individual needs
BYSC47N	school emphasizes sports
BYSC47O	students face competition for grades
BYSC48A	visitors required to sign in main office
BYSC48B	hall passes required to visit library
BYSC48D	hall passes required to visit office
BYSC48E	hall passes required to visit counselor
BYSC48F	academic counseling for students exists
BYSC48G	behavioral problem counseling for students exists
BYSC48H	vocational counseling for students exists
BYSC48I	student uniform required
BYSC48J	certain forms of dress forbidden
BYSC48K	students can't leave grounds during school hours
BYSC49A	degree student tardiness is a problem
BYSC49B	degree student absenteeism is a problem
BYSC49C	degree student class cutting is a problem
BYSC49D	degree student conflicts is a problem
BYSC49E	degree robbery or theft is a problem

variable	label
BYSC49F	degree vandalism is a problem
BYSC49G	degree student alcohol use is a problem
BYSC49H	degree student illegal drug use is a problem
BYSC49I	degree student weapons are a problem
BYSC49J	degree physical abuse of teachers is a problem
BYSC49K	degree verbal abuse of teachers is a problem
BYSC50AA	action for cheating: first occurrence
BYSC50AB	action for injury to other students: first occurrence
BYSC50AC	action for alcohol possession: first occurrence
BYSC50AD	action for drug possession: first occurrence
BYSC50AE	action for weapons possession: first occurrence
BYSC50AF	action for alcohol use: first occurrence
BYSC50AG	action for illegal drug use: first occurrence
BYSC50AH	action for smoking: first occurrence
BYSC50AI	action for verbal abuse of teachers: first occurrence
BYSC50AJ	action for injury to teacher: first occurrence
BYSC50AK	action for theft of school property: first occurrence
BYSC50AL	action for classroom disturbance: first occurrence
BYSC50AM	action for profanity: first occurrence
BYSC50BA	action for cheating: repeated occurrence
BYSC50BB	action for injury to other students: repeated occurrence
BYSC50BC	action for alcohol possession: repeat occurrence
BYSC50BD	action for drug possession: repeat occurrence
BYSC50BE	action for weapon possession: repeat occurrence
BYSC50BF	action for alcohol use: repeat occurrence
BYSC50BG	action for illegal drug use: repeat occurrence
BYSC50BH	action for smoking: repeat occurrence
BYSC50BI	action for verbal abuse of teacher: repeat occurrence
BYSC50BJ	action for injury to teacher: repeat occurrence
BYSC50BK	action for theft of school property: repeat occurrence
BYSC50BL	action for classroom disturbance: repeat occurrence
BYSC50BM	action for profanity: repeat occurrence

II. Analysis Plan

Maximum likelihood factor analysis was used, since it generally believed to be a better method than principal component, especially for large samples. Varimax rotation was applied, factor loadings greater than .4 were considered to be meaningful. After plausible grouping of the items was accomplished the composites were calculated.

Reliability of the suggested composites was evaluated by Cronbach's Alpha using SAS PROC CORR.

Validity was evaluated by regression R^2 using standardized math, history, reading and science scores (aggregated to school level) as the dependent variables and derived composites as the independent variables.

To compare the predictive power of the original item pool with the predictive power of the grouped item pool (composites and ungrouped items) each achievement score was regressed on:

a) The derived composites and the items not included in any of the composites.

b) All the original items.

III. Results

After appropriate items were reversed, all items were standardized (mean=0 and std=1).

FACTOR ANALYSIS

The original NELS:88 Base Year School File had 1035 observations. Following listwise deletion 966 observations were retained (missing values were evenly spread over all the items). All the calculations were performed using weights provided with the file. A total of 63 items were used.

Factor analyses with less than seven factors gave items loaded on multiple factors, therefore
seven, eight, and nine factor models were considered:

No. factors	proportion of variance explained	rms off-diagonal partials
7	.445	.069
8	.465	.065
9	.483	.061

Seven Factor Model

Rotated Factor Pattern (Varimax)

item	F1	F2	F3	F4	F5	F6	F7	label
BYSC50BG	.900	.007	.117	.044	.104	.154	.033	actn for illeg drug use: rep
BYSC50BF	.863	.030	.171	.056	.132	.179	.019	actn for alcohol use: rep
BYSC50BD	.867	.011	.068	.050	.157	.204	.102	action for drugs poss: rep
BYSC50BC	.793	.043	.132	.057	.205	.256	.003	action for alcohol poss: rep
BYSC50BE	.759	.027	.047	.004	.170	.170	.050	actn for weapon poss: rep
BYSC50BJ	.559	.038	.025	.045	.386	.091	-.031	action for injury to tch: rep
BYSC49K	.064	.672	.139	.127	.004	.024	-.007	verbal abuse teachers probl
BYSC49I	.000	.669	.046	.001	-.111	.049	.081	degree student weapon
BYSC49D	.011	.632	.160	.145	.013	.067	.115	degree student phys conflct
BYSC49E	.003	.630	.228	.099	-.001	.070	.137	degree robbery or theft
BYSC49F	.020	.615	.195	.182	-.042	.079	.074	degree vandalism problem
BYSC49J	.018	.611	.016	.008	-.068	-.007	.076	degree phys abuse of teach
BYSC49C	.005	.601	.256	.193	-.040	.047	.184	degree student class cutting
BYSC49B	.047	.556	.317	.171	.024	.050	.059	degree student absenteeism
BYSC49A	.039	.537	.267	.139	-.042	.023	.085	degree student tardiness
BYSC47H	.017	.308	.162	.282	-.098	.017	.101	teachr neg attitude to stdts
BYSC47I	.003	.311	.177	.156	-.089	.020	.017	difficulty motivating studs
BYSC48D	.088	.187	.886	.092	-.026	-.006	-.015	passes required to visit offc
BYSC48E	.096	.199	.854	.038	.000	.042	.048	passes reqrd to visit councl
BYSC48B	.098	.149	.821	.028	-.065	.017	-.013	passes required to vist libra
BYSC48C	.065	.200	.793	.067	-.052	-.039	-.042	passes required to vist lavat
BYSC48F	.049	.112	.421	.032	-.114	.156	.174	academic counseling exists
BYSC48H	.053	.234	.421	.032	.055	.058	.186	vocational counseling exists
BYSC48A	.029	.164	.381	.025	-.113	-.028	-.041	visitors reqd sign main off
BYSC48G	.027	.068	.301	.007	-.134	.162	.057	behavioral counseling exists
BYSC47N	.056	.086	.169	.165	-.023	-.049	-.142	school emphasize sports
BYSC48I	-.050	-.125	-.353	-.157	-.007	-.056	-.088	student uniform required
BYSC47E	.117	.143	.189	.811	-.058	-.021	.015	teachers encourage stdts
BYSC47F	.081	.100	.085	.767	-.048	.017	.074	stdt expctd to do homewrk
BYSC47J	.037	.042	.021	.640	-.031	.054	.062	school day is structured
BYSC47D	.009	.126	.149	.640	.038	.069	.105	clssrm environmt sructured
BYSC47G	.046	.196	.035	.621	.056	.096	.031	teacher moral is high
BYSC47B	.065	.002	.004	.601	.014	.038	.114	discipline is emphasized
BYSC47M	.047	.271	.166	.598	.071	.046	.094	tchrs respond to ind needs
BYSC47K	.011	.040	.046	.560	.003	.030	.105	rule deviation not tolerated

BYSC47C	.003	.246	.210	.448	.092	-.001	-.019	stdnts priority on learning
BYSC47L	.045	.095	.004	.257	.089	.039	.032	environment is " flexible"
BYSC47O	.107	.044	.017	.244	.039	-.043	-.058	studnts compete for grades
BYSC48J	.073	.031	.069	.169	-.121	.075	-.017	certain dress frms forbiddn
BYSC47A	.028	.200	.000	.329	.010	-.083	-.005	conflict: tchrs & administr
BYSC50BM	.132	.017	.003	.086	.673	.029	-.018	action for profanity: rep
BYSC50BK	.397	.116	.095	.106	.624	.135	-.080	action for theft: rep occ
BYSC50BL	.103	.057	.086	.032	.622	-.017	-.040	actn for cls disturbanc: rep
BYSC50BB	.300	.017	.031	.070	.601	.034	-.093	action for injury stud: rep
BYSC50BI	.332	.195	.108	.094	.577	.066	.012	actn for vrb abuse tch: rep
BYSC50BA	.123	.049	.039	.090	.511	.148	-.085	action for cheating: rep
BYSC50AK	.137	.037	.042	.012	.457	.218	.056	action for theft: 1st
BYSC50BH	.325	.229	.210	.169	.457	.173	.050	action for smoking: rep
BYSC50AI	.036	.003	.065	.035	.367	.054	.052	actn for vrb abuse tch: 1st
BYSC50AH	.031	.159	.103	.130	.357	.305	.070	action for smoking: 1st
BYSC50AM	.075	.178	.074	.009	.354	.064	.069	action for profanity: 1st occ
BYSC50AB	.010	.133	.240	.035	.350	.041	-.001	actn for injury to stud: 1st
BYSC50AJ	.315	.035	.101	.021	.340	.307	-.061	action for injury to tchr: 1st
BYSC50AL	.048	.206	.080	.032	.314	.120	.061	actn for cls distrubanc: 1st
BYSC50AA	.031	.019	.003	.012	.300	.186	-.024	action for cheating: 1st occ
BYSC50AG	.292	.813	.115	.057	.229	.855	-.015	actn for drug use: 1st occ
BYSC50AF	.238	.112	.140	.104	.237	.831	.029	action for alcohol use: 1st
BYSC50AD	.272	.046	.084	.061	.260	.824	-.015	actn for drug poss: 1st occ
BYSC50AC	.203	.069	.048	.108	.241	.738	.007	action for alcohol poss: 1st
BYSC50AE	.282	.022	.021	.010	.273	.585	.061	action for weapon poss: 1st
BYSC49G	.009	.396	.274	.150	.021	.131	.753	degree std alcohl use probl
BYSC49H	.030	.503	.281	.181	-.017	.090	.639	degree std drug use probl
BYSC48K	.046	.010	.006	.013	-.001	.053	-.208	std can't leave sch grounds

Following seven composites were identified:

Composite 1 SERIOUS REPEATED OFFENSES

variable	label
BYSC50BG	action for illegal drug use: repeat
BYSC50BF	action for alcohol use: repeat
BYSC50BD	action for drug possession: repeat
BYSC50BC	action for alcohol possession: repeat
BYSC50BE	action for weapon possession: repeat
BYSC50BJ	action for injury to teacher: repeat

Composite 2 STUDENT BEHAVIOR

variable	label
BYSC49K	verbal abuse teachers problem
BYSC491	degree student weapon problem
BYSC49D	degree student physical conflicts
BYSC49E	degree robbery or theft
BYSC49F	degree vandalism problem
BYSC49J	degree physical abuse of teacher
BYSC49C	degree student class cutting
BYSC49B	degree student absenteeism
BYSC49A	degree student tardiness
BYSC49H	degree student drug use problem

Composite 3 PASSES REQUIRED

variable	label
BYSC48D	passes required to visit office
BYSC48E	passes required to visit counselor
BYSC48B	passes required to visit library
BYSC48C	passes required to visit lavatory
BYSC48F	academic counseling exists
BYSC48H	vocational counseling exists

Composite 4 DISCIPLINE AND STRUCTURE

variable	label
BYSC47E	teachers encourage students
BYSC47F	student expected to do homework
BYSC47J	school day is structured
BYSC47D	classroom environment sructured
BYSC47G	teacher morale is high
BYSC47B	discipline is emphasized
BYSC47M	teacher respond to individual needs
BYSC47K	rule deviation not tolerated
BYSC47C	students priority on learning

Composite 5 MODERATE OFFENSES

variable	label
BYSC50BM	action for profanity: repeat
BYSC50BK	action for theft: repeat occurrence
BYSC50BL	action for class disturbance: repeat occurrence
BYSC50BB	action for injury student: repeat occurrence
BYSC50BI	action for verbal abuse teacher: repeat occurrence
BYSC50BA	action for cheating: repeat occurrence
BYSC50AK	action for theft: first occurrence
BYSC50BH	action for smoking: repeat occurrence

Composite 6 DRUG/ALCOHOL OFFENSES FIRST OCCURRENCE

variable	label
BYSC50AG	action for drug use: first occurrence
BYSC50AF	action for alcohol use: first occurrence
BYSC50AD	action for drug possession: first occurrence
BYSC50AC	action for alcohol possession: first occurrence
BYSC50AE	action for weapon possession: first occurrence

Composite 7 DEGREE DRUGS/ALCOHOL ARE THE PROBLEM

variable	label
BYSC49G	degree student alcohol use a problem
BYSC49H	degree student drug use a problem

RELIABILITY

Cronbach's Alpha was calculated for the following groups of items:

Composite 1	.925
Composite 2	.864
Composite 3	.861
Composite 4	.868
Composite 5	.825
Composite 6	.916
Composite 7	.874

Regressions were performed using standardized math, science and reading scores as a dependent variable and the seven composites as independent variables:

Dependent variable math score model R^2 =.012

F value(overall model)=1.669 P value=.113

The model was not significant at significance level=.05

Dependent variable science score model R^2 =.021

F value(overall model)=2.795 P value=.007

independent variable	parameter estimate	t-value	p-value
Intercept	50.55	304.21	.000
Composite1	.153	.604	.546
Composite2	.431	1.299	.194
Composite3	.704	3.001	.003
Composite4	.242	.961	.337
Composite5	-.441	-1.441	.15
Composite6	.182	.799	.424
Composite7	-.444	-1.96	.05

Dependent variable reading score model R^2 =.018

F value(overall model)=2.489 P value=.0155

The model was not significant at significance level=.05.

To compare the predictive power of the original item pool with the predictive power of the grouped item pool (composites and ungrouped items) each achievement score was regressed on:

a) The derived composites and the items not included in any of the composites.

b) All the original items.

All the models were significant at significance level=.05

dependent variable	composites + not grouped items model R^2	all items model R^2
math score	.063	.163
science score	.050	.150
reading score	.080	.219

IV. Analysis

The factor analyses demonstrated that eight and nine factor models did not give substantial increase in percent variance explained and there was no tangible decrease in root mean square off-diagonal partials. Moreover no new meaningful factors emerged in eight and nine factor models. This suggested that the seven factor model was the best:

Composite 1 SERIOUS REPEATED OFFENSES

variable	label
BYSC50BG	action for illegal drug use: repeat
BYSC50BF	action for alcohol use: repeat
BYSC50BD	action for drug possession: repeat
BYSC50BC	action for alcohol possession: repeat
BYSC50BE	action for weapon possession: repeat
BYSC50BJ	action for injury to teacher: repeat

Composite 2 STUDENT BEHAVIOR

variable	label
BYSC49K	verbal abuse teachers problem
BYSC491	degree student weapon problem
BYSC49D	degree student physical conflicts
BYSC49E	degree robbery or theft
BYSC49F	degree vandalism problem
BYSC49J	degree physical abuse of teacher
BYSC49C	degree student class cutting
BYSC49B	degree student absenteeism
BYSC49A	degree student tardiness
BYSC49H	degree student drug use problem

Composite 3 PASSES REQUIRED

variable	label
BYSC48D	passes required to visit office
BYSC48E	passes required to visit counselor
BYSC48B	passes required to visit library
BYSC48C	passes required to visit lavatory
BYSC48F	academic counseling exists
BYSC48H	vocational counseling exists

Composite 4 DISCIPLINE AND STRUCTURE

variable	label
BYSC47E	teachers encourage students
BYSC47F	student expected to do homework
BYSC47J	school day is structured
BYSC47D	classroom environment sructured
BYSC47G	teacher morale is high
BYSC47B	discipline is emphasized
BYSC47M	teacher respond to individual needs
BYSC47K	rule deviation not tolerated
BYSC47C	students priority on learning

Composite 5 MODERATE OFFENSES

variable	label
BYSC50BM	action for profanity: repeat
BYSC50BK	action for theft: repeat occurrence
BYSC50BL	action for class disturbance: repeat occurrence
BYSC50BB	action for injury student: repeat occurrence
BYSC50BI	action for verbal abuse teacher: repeat occurrence
BYSC50BA	action for cheating: repeat occurrence
BYSC50AK	action for theft: first occurrence
BYSC50BH	action for smoking: repeat occurrence

Composite 6 DRUG/ALCOHOL OFFENSES FIRST OCCURRENCE

variable	label
BYSC50AG	action for drug use: first occurrence
BYSC50AF	action for alcohol use: first occurrence
BYSC50AD	action for drug possession: first occurrence
BYSC50AC	action for alcohol possession: first occurrence
BYSC50AE	action for weapon possession: first occurrence

Composite 7 DEGREE DRUGS/ALCOHOL ARE THE PROBLEM

variable	label
BYSC49G	degree student alcohol use a problem
BYSC49H	degree student drug use a problem

With the remaining 18 variables not loading highly on any composite:

variable	label
BYSC47H	teacher negative attitude to students
BYSC47I	difficulty motivating students
BYSC48A	visitors required to sign in at main office
BYSC48G	behavioral counseling exists
BYSC47N	school emphasize sports
BYSC48I	student uniform required
BYSC47L	environment is "flexible"
BYSC47O	students compete for grades
BYSC48J	certain forms of dress forbidden
BYSC47A	conflict: teachers and administrators
BYSC50AI	action for verbal abuse teachers: first occurrence
BYSC50AH	action for smoking: first occurrence
BYSC50AM	action for profanity: first occurrence
BYSC50AB	action for injury to student: first occurrence
BYSC50AJ	action for injury to teacher: first occurrence
BYSC50AL	action for class disturbance: first occurrence
BYSC50AA	action for cheating: first occurrence
BYSC48K	students can't leave school grounds

Cronbach's Alpha (all alphas> .8) indicated that all the composites had high internal consistency. In other words, all the composites had a high percentage of their variation attributable to the "common score".

The validation results indicated that the composites had almost no predictive power. Therefore, while being internally consistent the composites had very weak validity as far as the achievement scores were concerned.

The composites were much weaker predictors of the achievement scores than the individual items.

V. Conclusion

The results of the analysis indicate that the 45 items from the school climate pool can be grouped into the seven composites with the remaining 18 items unattached. While each of the suggested composites demonstrated high internal consistency, grouping of the variables seriously reduces item pool predictive power thus making the composites useless. Therefore, forming composites from this item pool is not recommended

National Education Longitudinal Study: 1988 Base Year (NELS:88)

D. SCHOOL CLIMATE (Student Level)

I. Introduction

This section investigated possibilities of constructing school climate composites. The analysis was performed at the student level using the following 27 items from the NELS:88 Base Year Student File:

BYS57A	R had something stolen at school
BYS57B	someone offered to sell R drugs at school
BYS57C	someone threatened to hurt R at school
BYS58A	student tardiness a problem at school
BYS58B	student absenteeism a problem at school
BYS58C	student cutting class a problem at school
BYS58D	physical conflicts among stud a problem
BYS58E	robbery or theft a problem at school
BYS58F	vandalism of school property a problem
BYS58G	student use of alcohol a problem at school
BYS58H	student use of illegal drugs a problem
BYS58I	student possession of weapons a problem
BYS58J	physical abuse of teachers a problem
BYS58K	verbal abuse of teachers a problem
BYS59A	students get along well with teacher
BYS59B	there is real school spirit
BYS59C	rules for behavior are strict
BYS59D	discipline is fair
BYS59E	other students often disrupt class
BYS59F	the teaching is good
BYS59G	teachers are interested in students
BYS59H	teachers praise my effort
BYS59I	in class I feel put down by my teachers
BYS59J	most of my teachers listen to what I say
BYS59K	I don't feel safe at this school
BYS59L	student disruptions inhibit learning
BYS59M	misbehaving studs often get away with it

II. Analysis Plan

Maximum likelihood factor analysis was used to group the items. Varimax rotation was performed and loadings greater than .4 were considered to be meaningful. After plausible grouping of the items was accomplished the composites were calculated as the mean of the corresponding items.

Reliability of the suggested composites was evaluated by Cronbach's Alpha using SAS PROC CORR.

Validity was evaluated by using the R^2 values calculated from regressions of standardized math, science and reading scores on the derived composites.

To compare the composites with original items each achievement score was regressed on:

a) The derived composites and the items not included in any other composites.
b) All the original items.

III. Results

After the appropriate items were reversed, all items were standardized (mean=0 and std=1). NELS:88 Base Year Student File had a total of 24599 observations. Following listwise deletion 21642 observations were retained (missing values were evenly spread over the items).

Factor analysis (3 factor model)

Rotated Factor Pattern (Varimax).

Proportion variance explained=.364
RMS=.062

item	F1	F2	F3	label
BYS58H	0.848	0.895	0.135	student use of drugs
BYS58I	0.819	0.116	0.138	student possession of weapons
BYS58G	0.804	0.094	0.157	student use of alcohol
BYS58J	0.725	0.048	0.019	physical abuse of teachers
BYS58E	0.702	0.076	0.323	robbery or theft
BYS58F	0.693	0.071	0.292	vandalism of school property
BYS58K	0.693	0.146	0.210	verbal abuse of teachers
BYS58C	0.591	0.056	0.530	student cutting classes
BYS58D	0.554	0.078	0.496	physical conflict among students
BYS57B	0.186	0.181	0.023	someone offered to sell drugs
BYS59L	0.030	0.757	0.020	student disruptions learning
BYS59G	0.030	0.757	0.020	teachers interested in students
BYS59F	0.033	0.708	0.017	the teaching is good
BYS59J	0.028	0.678	0.005	teachers listen to what I say
BYS59H	-0.006	0.600	-0.014	teachers prize my effort
BYS59A	0.056	0.553	0.119	students/teachers get along
BYS59D	0.024	0.462	0.263	discipline is fair
BYS59I	0.075	0.438	-0.000	feel put down by my teachers
BYS59B	0.032	0.435	0.072	there is real school spirit
BYS59K	0.117	0.309	0.039	I don't feel safe at school
BYS59M	0.108	0.182	0.071	misbehaving stdts get away/w it
BYS57C	0.106	0.148	0.073	someone threaten to hurt
BYS57A	0.105	0.115	0.085	had something stolen in school
BYS59C	-0.024	0.029	0.027	rules for behavior are strict
BYS58A	0.232	0.048	0.723	student tardiness is a problem
BYS58B	0.290	0.032	0.723	student absenteeism
BYS59E	0.110	0.059	0.145	students often disrupt class
Proportion variance explained	0.179	0.114	0.072	

Factor analysis (four factor model)

Proportion variance explained=.404
RMS=.049

item	F1	F2	F3	F4	label
BYS58H	0.865	0.067	0.042	0.053	student use of drugs
BYS58G	0.827	0.070	0.071	0.029	student use of alcohol
BYS58I	0.819	0.073	0.040	0.154	weapon possession
BYS58E	0.719	0.035	0.233	0.172	robbery or theft
BYS58J	0.708	0.031	0.205	0.162	physical abuse of teachers
BYS58F	0.708	0.031	0.205	0.162	vandalism of school proper
BYS58C	0.640	0.031	0.460	0.109	stds cutting classes
BYS58K	0.622	0.112	0.135	0.135	verbal abuse of teachr
BYS58D	0.597	0.048	0.425	0.141	physical abuse
BYS57B	0.188	0.166	0.007	0.084	offered to sell drugs
BYS59G	0.060	0.760	0.015	0.043	tchrs interested in stds
BYS59F	0.059	0.706	0.010	0.056	the teaching is good
BYS59J	0.052	0.678	0.000	0.048	teachers listen to me
BYS59H	0.019	0.610	0.009	0.009	teachers prize my effort
BYS59A	0.077	0.536	0.098	0.138	studs/teachers get along
BYS59D	0.043	0.460	0.022	0.031	discipline is fair
BYS59B	0.045	0.421	0.055	0.117	there is school spirit
BYS59I	0.064	0.407	0.041	0.254	feel put down by tchrs
BYS59C	0.018	0.032	0.031	0.008	behavior rules- strict
BYS58A	0.307	0.038	0.692	0.089	student tardiness
BYS58B	0.362	0.019	0.686	0.097	student absenteeism
BYS59L	0.306	-0.085	0.003	0.606	stdts disrupt learning
BYS59M	0.063	0.111	-0.007	0.515	misbehaving std get away
BYS59E	0.07	0.005	0.082	0.427	students disrupt class
BYS59K	0.089	0.258	-0.024	0.380	don't feel safe school
BYS57C	0.087	0.106	0.024	0.278	someone threatened
BYS57A	0.090	0.077	0.042	0.245	had something stolen
Proportion variance explained	0.189	0.108	0.055	0.051	

Following four composites were identified:

Composite 1

BYS58H	student use of illegal drugs a problem
BYS58G	student use of alcohol a problem
BYS58I	student possession of weapons a problem
BYS58E	robbery or theft problem at school
BYS58J	physical abuse of teachers a problem
BYS58F	vandalism of school property a problem
BYS58C	students cutting classes a problem
BYS58K	verbal abuse of teachers a problem
BYS58D	physical conflicts among students a problem

Composite2

BYS59G	teachers are interested in students
BYS59F	the teaching is good
BYS59J	most of my teachers listen to what I say
BYS59H	teachers praise my effort
BYS59A	students get along well with teachers
BYS59D	discipline is fair
BYS59B	there is real school spirit
BYS59I	in class I feel put down by my teachers

Composite 3

BYS58A	student tardiness is a problem
BYS58B	student absenteeism is a problem

Composite 4

BYS59L	student disruption inhibit learning
BYS59M	misbehaving student often get away with it
BYS59E	other students often disrupt class

Reliability

Four factor solution.

Cronbach's Alphas were calculated for the following groups of items:

Composite 1 .922

Composite 2 .801

Composite 3 .772

Composite 4 .564

Regressions were performed using standardized math, science and reading scores as dependent variables and the four composites as the independent variables:

Dependent variable math score model R^2=.018

F-value(overall)=97.5 p-value=.0001

independent variable	parameter estimate	t-value	p-value
Intercept	50.628	741.635	.00
Composite1	3.233	2.822	.005
Composite2	-15.882	-13.269	.0001
Composite3	2.037	2.064	.039
Composite4	-14.213	-13.435	.0001

Dependent variable science score model R^2=.015

F-value(overall)=79.17 p-value=.0001

independent variable	parameter estimate	t-value	p-value
Intercept	50.60	744.213	.00
Composite1	.868	.761	.448
Composite2	-15.226	-12.778	.0001
Composite3	1.027	1.045	.296
Composite4	-11.616	-11.019	.0001

Dependent variable reading score model R^2=.023

F-value(overall)=122.918 p-value=.0001

independent variable	parameter estimate	t-value	p-value
Intercept	50.64	751.021	.00
Composite1	.647	.572	.567
Composite2	-16.53	-13.972	.0001
Composite3	7.669	7.869	.0001
Composite4	-15.164	-14.512	.0001

Regressions were performed using the achievement scores as dependent variables and

a) The four composites and the not grouped items.
b) All the items.

All models were significant at significance level=.05

dependent variable	composites + not grouped items model R^2	all items model R^2
math score	.047	.123
science score	.044	.118
reading score	.059	.141

IV. Analysis

Factor analysis: three, four, five, and six factor models were run. The four factor model had four factors consistently substantiated: five and six factor models had the same four factors as the four factor model plus factors with no strong loadings. Thus, the four factor model (with five items not a part of any composite) was chosen.

In addition, throughout the analysis items BYS58C and BYS58D had strong loadings on one factor and higher than .4 loadings on the third factor. Addition of these two items did not lead to a meaningful increase in reliability of the third factor therefore these items were used in the first factor only.

Cronbach's Alpha indicated good internal consistency for the first three composites, while the fourth composite had rather low reliability.

The validation results demonstrated that the composites had almost no predictive power. The composites were much weaker predictors of the achievement scores than the individual items.

V. Conclusion

The analysis suggested that the 27 item pool of student level school climate could be grouped into four composites and five unattached items:

Composite 1 Student violence/misbehavior

Composite 2 Perception of Teachers/ school spirit

Composite 3 Student tardiness/absenteeism

Composite 4 disruption of learning by students

At the same time the grouping dramatically reduces the item pool predictive power, thus making the composites, not as valuable predictors as individual items.

E. VALIDATION OF THE SET OF NELS:88 COMPOSITES

Predictive power of the composites identified as most effective, and the item pools was tested on student and school levels in the following way:

a) Student level: math, reading and science scores were regressed on the SES, locus of control, self concept and all the 27 items from the student level school climate pool. School climate individual items were used instead of composites since it was shown that school climate composites had weak predictive power. Composites selected in previous sections as the most effective were used, i.e. SES2, LC3 and SC2.

b) School level: aggregated math, reading and sciences scores were regressed on the aggregated ses, locus of control, self concept composites and 63 items from the school level school climate pool. School climate individual items were used since it was shown that school climate composites had almost no predictive power. Composites selected in previous sections as the most effective were used, i.e. SES2, LC3 and SC2.

Student Level Results

Stepwise regressions were performed using significance level for staying=.05 and significance level of entry =.05. The following variables were selected:

Dependent variable math score (n=19510)

step	variable	model R^2	label
1	SES2	.187	socio-economic status
2	LC3	.234	locus of control
3	BYS59L	.241	student disruption inhibits learning
4	BYS58K	.247	verbal abuse of teachers is problem
5	BYS58J	.255	physical abuse of teachers is probl
6	BYS57B	.260	someone offered to sell drugs
7	BYS59A	.264	students get along well with teachers
8	BYS59M	.267	misbehaving students get away
9	BYS58E	.269	robbery and theft is a problem
10	BYS58I	.271	possession of weapons is a problem
11	BYS58G	.273	use of alcohol is problem
12	BYS59K	.275	I don't feel safe at school
13	BYS59C	.276	rules for behavior are strict
14	BYS57A	.278	student had something stolen
15	BYS57C	.278	someone threaten to hurt at school

Another 10 variables were selected at significance level=.05 but had partial $R^2<.001$. The final model for 25 variables had an $R^2=.283$

Dependent variable science score (n=19494)

step	variable	model R^2	label
1	SES2	.145	socio-economic status
2	LC3	.188	locus of control
3	BYS59L	.194	student disruption inhibits learning
4	BYS58K	.198	verbal abuse of teachers is problem
5	BYS58J	.212	physical abuse of teachers is probl
6	BYS59A	.217	students get along well with teachers
7	BYS57B	.220	someone offered to sell drugs
8	BYS59M	.223	misbehaving students get away
9	BYS59K	.225	I don't feel safe at school
10	BYS57C	.227	someone threaten to hurt at school
11	BYS58E	.229	robbery and theft is a problem
12	BYS59C	.230	rules for behavior are strict
13	BYS58I	.231	possession of weapons is a problem
14	BYS58G	.232	use of alcohol is problem
15	BYS57A	.233	student had something stolen
16	BYS58C	.234	student cutting classes a problem
17	BYS58D	.235	physical conflict among students
18	BYS59H	.236	teachers praise my effort
19	BYS59F	.237	the teaching is good

Another five variables were selected at significance level=.05 but had partial $R^2<.001$
The final model for 24 variables had an $R^2=.239$

Dependent variable reading score (n=19516)

step	variable	model R^2	label
1	SES2	.148	socio-economic status
2	LC3	.207	locus of control
3	BYS59L	.215	student disruption inhibits learning
4	BYS58K	.223	verbal abuse of teachers is problem
5	BYS58J	.240	physical abuse of teachers is probl
7	BYS57B	.245	someone offered to sell drugs
8	BYS59K	.249	I don't feel safe at school
9	BYS58E	.253	robbery and theft is a problem
10	BYS59A	.255	students get along well with teachers
11	SC2	.259	self concept
12	BYS57A	.261	student had something stolen
13	BYS58I	.264	possession of weapons is a problem
14	BYS58B	.266	student absenteeism is a problem
15	BYS59C	.267	rules for behavior are strict
16	BYS59F	.268	the teaching is good
17	BYS58G	.269	use of alcohol is problem

Another eight variables were selected at significance level=.05 but had partial $R^2<.001$. The final model for 25 variable had an $R^2=.272$

School Level Results

Stepwise regressions were performed using significance level for staying=.05 and significance level of entry =.05. The following variables were selected:

Dependent variable math score (n=948)

step	variable	model R^2	label
1	SES2	.651	socio-economic status
2	LC3	.684	locus of control
3	SC2	.697	self concept
4	BYSC47C	.703	students place priority on learning
5	BYSC48A	.707	visitors required to sign in
6	BYSC49D	.710	degree student phys conflict prob
7	BYSC49G	.713	degree student alcohol problem
8	BYSC50BM	.716	action for profanity: repeat occur
9	BYSC49I	.718	degree student weapons problem
10	BYSC47B	.719	discipline is emphasized
11	BYSC47O	.721	students compete for grades
12	BYSC47L	.722	school environment is flexible
13	BYSC50AC	.723	action for alcohol poss: 1st occur

Other variables were not entered at significance level=.05

Dependent variable science score (n=947)

step	variable	model R^2	label
1	SES2	.538	socio-economic status
2	LC3	.582	locus od control
3	SC2	.598	self concept
4	BYSC49I	.604	degree student weapons problem
5	BYSC47O	.608	student compete for grades
6	BYSC48A	.612	victors required to sign in
7	BYSC49G	.616	degree student alcohol problem
8	BYSC50BM	.619	action for profanity: rep occur.
9	BYSC49C	.622	degree class cutting is a prob
10	BYSC47C	.624	students place priority on learning
11	BYSC48I	.626	student uniform required
12	BYSC50AB	.627	action for injury to other stud: 1st

Other variables were not entered at significance level=.05

Dependent variable reading score (n=947)

step	variable	model R^2	label
1	SES2	.615	socio-economic status
2	LC3	.673	locus of control
3	SC2	.693	self concept
4	BYSC48B	.703	hall passes required for library
5	BYSC47C	.708	students place priority on learning
6	BYSC49A	.711	degree student tardiness a problem
7	BYSC48A	.713	victors required to sign in
8	BYSC49D	.715	degree student phys conflict prob
9	BYSC47L	.716	school environment is flexible
10	BYSC50AB	.718	action for injury to other students: 1st

Other variables were not entered at significance level=.05

Analysis

Both student and school level analyses demonstrated that after correcting for variation in achievement scores due to SES remaining variables added only moderately to the model's predictive power. SES always enters the model first. Partial R^2s for locus of control, self concept and other variables were much lower when they were in the model with SES than they were when SES was not in the model. For example, math score (student level) regressed on LC3 alone had R^2=.103, while the partial R^2 with SES2 in the model was .043 . This was especially apparent at the school level.

Therefore, as far as predictive power of the composites and the items is concerned, SES dominates and seemed to cover most of the predictive power of the combined composite/item pool.

OVERALL SUMMARY

Three potential composite areas were considered in the analysis of NELS data:

1) Locus of control/self concept item pool.
2) SES item pool.
3) School Climate item pool.

Only the SES composite was capable of condensing the information of the original items while preserving the predictive power. SES was also by far the most powerful predictor of the achievement scores, especially at the school level.

Locus of control composite did lose about 15% of its R^2 compared to individual items but was still a relatively good predictor of achievement scores. The rest of the composites were of less use as predictors.

Even as individual items, school climate variables were ineffective predictors, after SES and locus of control were entered in the models:

a) on the student, level significant school climate items contributed less then 25% of the combined model R^2

b) on the school level, significant school climate items contributed 5-7% to the combined model R^2

SCHOOL CLIMATE
SASS

I. Introduction

This section investigates possibilities of constructing school climate composites. The analysis was done on the teacher level using 36 items from the 1987-1988 School and Staffing Survey Teacher File:

TSC238	teachers are evaluated fairly
TSC239	principal lets staff know what is expected
TSC240	administ behavior is supportive and encouraging
TSC241	I am satisfied with my teaching salary
TSC242	misbehavior interferes with my teaching
TSC243	teachers participate in important decisions
TSC244	receive parental support for work
TSC245	necessary materials are available
TSC246	principal does poor job of getting resources
TSC247	routine duties/paperwork interferes w/teaching
TSC248	my principal enforces rules for conduct
TSC249	principal talks w/me about instructional practices
TSC250	student behavior rules enforced by all teachers
TSC251	colleagues share my belief/values about school
TSC252	principal knows school goals and communicates
TSC253	there is great cooperation among staff
TSC254	staff members recognized for job well done
TSC255	follow rules that conflict w/my judgement
TSC256	I am satisfied with my class sizes
TSC257	make effort to coordinate course students
TSC258	goals/priorities for school are clear
TSC259	student tardiness/class cut interferes w/teaching
TSC260	it is waste of time to do my best as teacher
TSC262	degree of problem- student tardiness
TSC263	degree of problem- student absenteeism
TSC264	degree of problem- teacher absenteeism
TSC265	degree of problem- student cutting classes
TSC266	degree of problem- physical conflict: students
TSC267	degree of problem- robbery or theft
TSC268	degree of problem- vandalism of school property
TSC269	degree of problem- student pregnancy
TSC270	degree of problem- student use of alcohol
TSC271	degree of problem- student drug abuse
TSC272	degree of problem- student possess weapons
TSC273	degree of problem- physical abuse of teachers

TSC274 degree of problem- verbal abuse of teachers

II. Analysis Plan

Maximum likelihood factor analysis method was used to group the items. Varimax rotation was performed, with loadings greater than .4 considered meaningful. After plausible grouping of the items was accomplished, the composites were calculated as the means of the selected items. Factor analysis was performed separately for public and private schools. Reliability of the suggested composites was evaluated by Cronbach's Alphas using SAS PROC CORR.

III. Results

The original SASS Teacher File had 47,537 observations. Following listwise deletion 43,397 observations were retained.

Factor analysis (five factor model)

Proportion of variance explained=.444
RMS=.044

Rotated Factor Pattern (varimax).

item	F1	F2	F3	F4	F5	label
TSC239	.790	.022	.033	.036	.028	principal lets staff know what is expected
TSC252	.752	.048	.045	.059	.182	principal knows schl goals & communicates
TSC240	.734	.115	.044	.027	.049	admins behavior is supportive and encour
TSC248	.676	.162	.008	.102	.043	my principal enforces rules for conduct
TSC254	.669	.095	.089	.028	.249	staff member recogn for job well done
TSC258	.642	.099	.064	.091	.300	goals/priorities for school are clear
TSC238	.629	.125	.050	.052	.100	teachers are evaluated fairly
TSC246	.586	.131	.018	.070	.050	princpl does poor job of getting resources
TSC249	.579	.020	.082	.060	.127	princpl talks with me about instr practices
TSC243	.509	.127	.111	.081	.224	teachers participate in important decisions

TSC255	.411	.226	.037	.104	.148	follow rules that conflict with my judgment
TSC245	.294	.231	-.018	.073	.149	necessary materials are available
TSC260	.288	.235	.079	.169	.182	it is waste of time to do best as teacher
TSC241	.145	.095	-.001	-.007	.065	I am satisfied with my teaching salary
TSC266	.106	.684	.063	.184	.056	degree of problem-physical conflict: students
TSC274	.191	.665	.243	.177	.099	degree of problem- verbal abuse of teachers
TSC267	.093	.636	.256	.206	.079	degree of problem- robbery or theft
TSC273	.095	.626	.162	.034	-.013	degree of problem- physical abuse of teachers
TSC272	.069	.620	.367	.117	.027	degree of problem- student poss of weapons
TSC268	.126	.612	.250	.227	.099	degree of problem- vandalism of school property
TSC242	.212	.437	-.004	.215	.143	misbehavior interferes w/my teaching
TSC264	.086	.363	.124	.309	.187	degree of problem- teacher absenteeism
TSC244	.190	.264	.093	.171	.251	receive parental support for work
TSC247	.179	.221	.023	.096	.096	routine duties/paperwrk interfere w/teaching
TSC256	.085	.180	-.063	.060	.099	I am satisfied with my class sizes
TSC270	.115	.176	.877	.218	.137	degree of problem- student use of alcohol
TSC271	.107	.294	.839	.237	.122	degree of problem- student drug abuse
TSC269	.062	.316	.642	.252	.100	degree of problem- student pregnancy
TSC262	.090	.331	.234	.737	.089	degree of problem- student tardiness
TSC263	.090	.351	.274	.687	.126	degree of problem- student absenteeism
TSC259	.150	.278	.199	.619	.119	stu tardiness/class cut interfere w/teaching
TSC265	.103	.385	.438	.516	.102	degree of problem-student cutting classes
TSC251	.208	.099	.084	.073	.623	colleagues share belief/value of school
TSC253	.375	.152	.069	.074	.579	there is great cooperation among staff
TSC250	.380	.197	.178	.215	.428	student behavior rules enforced by all teach
TSC257	.111	.029	.042	.039	.276	make effort to coordinate course content
Proportion of variance explained	.149	.110	.075	.065	.044	

The following 5 composites were identified:

Composite1	(teachers satisfaction with administrative support and leadership)
TSC239	principal lets staff know what is expected
TSC252	principal knows school goals and communicates
TSC240	admins behavior is supportive and encouraging
TSC248	my principal enforces rules for conduct
TSC254	staff members recognized for job well done
TSC258	goals/priorities for school are clear
TSC238	teachers are evaluated fairly
TSC246	principal does poor job of getting resources
TSC249	principal talks w/me about instructional practices
TSC243	teachers participate in important decisions
TSC255	have to follow rules that conflict w/my judgment
Composite2	(behavioral problems: violence)
TSC266	degree of problem- physical conflict students
TSC274	degree of problem- verbal abuse of teachers
TSC267	degree of problem- robbery or theft
TSC273	degree of problem- physical abuse of teachers
TSC272	degree of problem- student possess weapons
TSC268	degree of problem- vandalism of school property
TSC242	misbehaving in school interferes w/my teaching
Composite3	(behavior problems: substance abuse and pregnancy)
TSC270	degree of problem- student use of alcohol
TSC271	degree of problem- student drug abuse
TSC269	degree of problem- student pregnancy
TSC265	degree of problem- students cutting classes
Composite4	(behavior problems: absenteeism-tardiness)
TSC262	degree of problem- student tardiness
TSC263	degree of problem- student absenteeism
TSC259	tardiness/class cutting interferes w/teaching
TSC265	degree of problem- students cutting class
Composite5	(staff cooperation)
TSC251	colleagues share my belief/value for school
TSC253	there is great cooperation among staff
TSC250	rules for behavior enforced by all teachers

Reliability

Composite1	.895
Composite2	.855
Composite3	.894
Composite4	.869
Composite5	.712

IV. Analysis

The four-factor model did not distinguish between the "administration support and leadership" composite (composite1) and "cooperation among staff" composites. The five-factor model split those two composites. There was no new useful composites in the six-factor model and TSC265 (cutting classes) became loaded on three factors (it was loaded on two factors in the four- and five-factor models). Therefore, the five-factor model was selected. The following composites were formed:

Composite1	(teachers satisfaction with administrative support and leadership)
TSC239	principal lets staff know what is expected
TSC252	principal knows school goals and communicates
TSC240	admins behavior is supportive and encouraging
TSC248	my principal enforces rules for conduct
TSC254	staff members recognized for job well done
TSC258	goals/priorities for school are clear
TSC238	teachers are evaluated fairly
TSC246	principal does poor job of getting resources
TSC249	principal talks w/me about instructional practices
TSC243	teachers participate in important decisions
TSC255	have to follow rules that conflict w/my judgment

Composite2 (behavioral problems: violence)

TSC266	degree of problem- physical conflict students
TSC274	degree of problem- verbal abuse of teachers
TSC267	degree of problem- robbery or theft
TSC273	degree of problem- physical abuse of teachers
TSC272	degree of problem- student possess weapons
TSC268	degree of problem- vandalism of school property
TSC242	misbehaving in school interferes w/my teaching

Composite3 (behavior problems: substance abuse and pregnancy)

TSC270	degree of problem- student use of alcohol
TSC271	degree of problem- student drug abuse
TSC269	degree of problem- student pregnancy
TSC265	degree of problem- students cutting classes

Composite4 (behavior problems: absenteeism-tardiness)

TSC262	degree of problem- student tardiness
TSC263	degree of problem- student absenteeism
TSC259	tardiness/class cutting interferes w/teaching
TSC265	degree of problem- students cutting class

Composite5 (staff cooperation)

TSC251	colleagues share my belief/value for school
TSC253	there is great cooperation among staff
TSC250	rules for behavior enforced by all teachers

The remaining eight items did not have meaningful loadings on any of the factors.

Separate factor analyses for private and public school subsamples gave essentially the same
results, confirming the above grouping.

The first four composites had high Cronbach's Alphas, indicating high internal consistency. the fifth composite had adequate reliability.

V. Conclusion

The items from the teacher school climate pool could be grouped into following composites:

1) Composite1 - administrative support and leadership
2) Composite2 - behavioral problems (violent)
3) Composite3 - behavioral problems (drug abuse/pregnancy)
4) Composite4 - behavioral problems (absenteeism/tardiness)
5) Composite5 - staff cooperation

with remaining eight items not being part of any composite.

PERCEPTIONS OF SCHOOL PROBLEMS
SASS School Administrator Questionnaire

I. Introduction

This analysis explores the possibilities of forming composites from the "administrator's perception of the school problems" item pool. The following 13 items from the SASS School Administrator questionnaire were used:

item name	item label
ASC087	student tardiness
ASC088	student absenteeism
ASC089	teacher absenteeism
ASC090	student cutting class
ASC091	physical conflict among students
ASC092	robbery or theft
ASC093	vandalism of school property
ASC094	student pregnancy
ASC095	student use of alcohol
ASC096	student drug abuse
ASC097	student possession of weapons
ASC098	physical abuse of teachers
ASC099	verbal abuse of teachers

II. Analysis Plan

Maximum likelihood factor analysis was used to identify composites. After varimax rotation was performed, loadings greater than .4 were considered to be meaningful. Following identification of the items composites were computed as mean of the selected items.

Separate factor analysis was performed on public and private school subsamples.

Reliability of the suggested composites was evaluated by Cronbach's Alpha using SAS PROC CORR.

III. Results

The original SASS Administrator file contained 10955 records. Following listwise deletion, 10702 records were retained. All items were standardized (mean=0 and std=1).

Factor analysis (two-factor model)

Proportion variance explained=.462
RMS=.09

Rotated Factor Pattern (Varimax)

items	F1	F2	label
ASC091	.640	.033	physical conflicts
ASC099	.597	.232	verbal abuse of teachers
ASC097	.563	.285	student possession of weapons
ASC088	.559	.367	student absenteeism
ASC093	.551	.228	vandalism of school property
ASC092	.542	.346	robbery or theft
ASC089	.498	.178	teacher absenteeism
ASC087	.488	.355	student tardiness
ASC098	.443	.080	physical abuse of teachers
ASC095	.108	.917	student use of alcohol
ASC096	.234	.884	student drug use
ASC094	.306	.663	student pregnancy
ASC090	.440	.563	student cutting classes
Proportion of variance explained	.233	.230	

Factor analysis (three-factor model)

Proportion variance explained=.518
RMS=.052

Rotated Factor Pattern (Varimax)

items	F1	F2	F3	label
ASC095	.902	.109	.161	student use of alcohol
ASC096	.867	.222	.212	student drug use

ASC094	.625	.205	.321	student pregnancy
ASC090	.507	.257	.460	student cutting classes
ASC099	.202	.641	.194	verbal abuse of teachers
ASC091	.002	.605	.255	physical conflicts
ASC097	.255	.575	.211	student possession of weapons
ASC098	.071	.557	.026	physical abuse of teachers
ASC093	.196	.485	.271	vandalism of school property
ASC092	.314	.480	.278	robbery or theft
ASC088	.262	.234	.746	student absenteeism
ASC087	.264	.178	.668	student tardiness
ASC089	.118	.309	.441	teacher absenteeism
Proportion of variance explained	.201	.172	.145	

Factor analysis (four-factor model)

Proportion variance explained=.552
RMS=.033

Rotated Factor Pattern (Varimax)

items	F1	F2	F3	F4	label
ASC095	.899	.147	.061	.634	student use of alcohol
ASC096	.862	.207	.176	.166	student drug use
ASC094	.624	.321	.183	.123	student pregnancy
ASC090	.503	.450	.194	.201	student cutting classes
ASC088	.259	.757	.168	.173	student absenteeism
ASC087	.263	.653	.101	.187	student tardiness
ASC089	.114	.437	.235	.208	teacher absenteeism
ASC098	.069	.048	.645	.084	physical abuse of teachers
ASC099	.195	.212	.575	.271	verbal abuse of teachers
ASC097	.249	.227	.556	.217	student possession weapons
ASC091	.015	.249	.437	.404	physical conflicts
ASC092	.283	.215	.219	.634	robbery or theft
ASC093	.165	.217	.237	.596	vandalism of school property
Proportion of variance explained	.197	.140	.119	.097	

Reliability
(For the three factor solution)

Composite1:	ASC095	.877
	ASC096	
	ASC094	
	ASC090	
Composite2	ASC099	.786
	ASC091	
	ASC097	
	ASC098	
	ASC093	
	ASC092	
Composite3	ASC088	.776
	ASC087	
	ASC089	
	ASC090	

IV. Analysis

The factor analysis results demonstrated noticeable improvement in three-factor model over the two-factor model. In both models ASC90 is loaded on two factors. In the four-factor model ASC90 and ASC91 were loaded on more than one factor. In addition, the four-factor model did have slightly higher proportion of variance explained and lower RMS but the improvement four over three factor model is much less than the three-over two-factor models. Four factor model had a separate 'robbery/theft and vandalism factor'. However, the three factor model was selected.

Composite1

ASC095	student use of alcohol
ASC096	student drug abuse
ASC094	student pregnancy
ASC090	student cutting class

Composite2

ASC099	verbal abuse of teachers
ASC091	physical conflict among students
ASC097	student possession of weapons
ASC098	physical abuse of teachers
ASC093	vandalism of school property
ASC092	robbery or theft

Composite3

ASC088	student absenteeism
ASC087	student tardiness
ASC090	student cutting class
ASC089	teacher absenteeism

Separate factor analysis for private and public schools gave essentially the same results.

The Cronbach's Alphas indicated adequate internal consistency of the composites.

V. Conclusion

The "School Problem" items from the SASS Administrator File form three composites:

Composite1: substance abuse and pregnancy

Composite2: violence

Composite3: tardiness/absenteeism

Listing of NCES Working Papers to Date

Number	Title	Contact
94-01	Schools and Staffing Survey (SASS) Papers Presented at Meetings of the American Statistical Association	Dan Kasprzyk
94-02	Generalized Variance Estimate for Schools and Staffing Survey (SASS)	Dan Kasprzyk
94-03	1991 Schools and Staffing Survey (SASS) Reinterview Response Variance Report	Dan Kasprzyk
94-04	The Accuracy of Teachers' Self-reports on their Postsecondary Education: Teacher Transcript Study, Schools and Staffing Survey	Dan Kasprzyk
94-05	Cost-of-Education Differentials Across the States	William Fowler
94-06	Six Papers on Teachers from the 1990-91 SASS and Other Related Surveys	Dan Kasprzyk
94-07	Data Comparability and Public Policy: New Interest in Public Library Data Papers Presented at Meetings of the American Statistical Association	Carrol Kindel
95-01	Schools and Staffing Survey: 1994 papers presented at the 1994 Meeting of the American Statistical Association	Dan Kasprzyk
95-02	QED Estimates of the 1990-91 Schools and Staffing Survey: Deriving and Comparing QED School Estimates with CCD Estimates	Dan Kasprzyk
95-03	Schools and Staffing Survey: 1990-91 SASS Cross-Questionnaire Analysis	Dan Kasprzyk

Listing of NCES Working Papers to Date (Continued)

Number	Title	Contact
95-04	National Education Longitudinal Study of 1988: Second Follow-up Questionnaire Content Areas and Research Issues	Jeffrey Owings
95-05	National Education Longitudinal Study of 1988: Conducting Trend Analyses of NLS-72, HS&B, and NELS:88 Seniors	Jeffrey Owings
95-06	National Education Longitudinal Study of 1988: Conducting Cross-Cohort Comparisons Using HS&B, NAEP, and NELS:88 Academic Transcript Data	Jeffrey Owings
95-07	National Education Longitudinal Study of 1988: Conducting Trend Analyses HS&B and NELS:88 Sophomore Cohort Dropouts	Jeffrey Owings
95-08	CCD Adjustments to the 1990-91 SASS: A Comparison of Estimates	Dan Kasprzyk
95-09	The Results of the 1993 Teacher List Validation Study (TLVS)	Dan Kasprzyk
95-10	The Results of the 1991-92 Teacher Follow-up Survey (TFS) Reinterview and Extensive Reconciliation	Dan Kasprzyk
95-11	Measuring Instruction, Curriculum Content, and Instructional Resources: The Status of Recent Work	Sharon Bobbitt & John Ralph
95-12	Rural Education Data User's Guide	Samuel Peng

Listing of NCES Working Papers to Date (Continued)

Number	Title	Contact
95-13	Assessing Students with Disabilities and Limited English Proficiency	James Houser
95-14	Empirical Evaluation of Social, Psychological, & Educational Construct Variables Used in NCES Surveys	Samuel Peng